Method for Calculating the Residual Resource of Fog Node Elements of Distributed Information Systems of Critical Infrastructure Facilities

Anton Zahynei1, Yurii Shcheblanin2, Oleg Kurchenko2, Andrii Anosov3, and Vladyslav Kruglyk4

1 State University of Information and Communication Technologies, 7 Solomyanska str., Kyiv, 03110, Ukraine
2 Taras Shevchenko National University of Kyiv, 60 Volodymyrska str., Kyiv, 01033, Ukraine
3 Borys Grinchenko Kyiv Metropolitan University, 18/2 Bulvaro-Kudriavska str., Kyiv, 04053, Ukraine
4 Bogdan Khmelnitsky Melitopol state pedagogical university, 59 Naukovogo mistechka, Zaporizhzhya, 69000 Ukraine

Abstract
The paper considers the process of predicting the residual technical resource of elements (devices) of information and communication systems, which are components of fog computing. The analysis of the influence of the technical state of the elements of fog nodes on the functioning of distributed information systems is carried out. A method for calculating the residual resource of fog nodes has been developed, which is advisable to use for planning the modernization of fog nodes and in the process of calculating the fault tolerance of the entire distributed infrastructure.

Keywords
Fog computing, information systems, model, residual resource.

1. Introduction

Modern technology development and implementation of distributed information systems in critical areas of human activity necessitate the improvement of methods for forecasting and optimizing the use of technical resources. One of the key innovative technologies that has become an important component in modern information and communication systems is fog computing technology [1].

The fog node is the core component of the fog computing architecture. Fog nodes are either physical components (e.g. gateways, switches, routers, servers, etc.) or virtual components (e.g. virtualized switches, virtual machines, cloudlets, etc.) that are tightly coupled with the smart end-devices or access networks, and provide computing resources to these devices [2].

In distributed information systems of critical infrastructure, such as energy, transportation, electronic communications, and others, where security and uninterrupted operation are a priority, the use of fog nodes becomes an important strategic solution [3]. Node elements allow the distributed system to be more flexible and adaptable to the challenges it faces.

Fog nodes consist of devices that cannot process data on their own. Examples of such devices include temperature sensors, humidity sensors, motion sensors, smoke detectors, and video cameras without image processing capabilities.

Given the importance and unique capabilities of fog nodes in the context of critical infrastructure, there is a need to develop a method for calculating the remaining life of node elements [4–6]. This method will allow us to effectively determine and predict the resources remaining after devices have...
performed computing and networking tasks over a long period.

The research aims to develop and optimize fog computing technology for critical infrastructure objects. It will consider the residual resources of node elements to create more sustainable and efficient distributed information systems. The research will also consider failure rate criteria [7].

2. Review of Existing Methods

Most of the research related to the resource of fog node elements is to calculate their required amount to perform the tasks. Therefore, in [8] the authors present a review of resource reservation methods in a fog environment, i.e. the calculation of the amount of resources that will be sufficient for the stable functioning of the information system is investigated. In addition, the paper describes the methods that should be used to calculate the time to failure of fog node elements, such methods include a task execution model, statistical method, etc. [8]. The task execution model is used to calculate the execution time of a specific task by a system element. The model takes into account such factors as the type of task, the size of the input data, and the available resources of the tool [8]. In turn, the statistical method uses statistical data to calculate the execution time of the task. The method takes into account such criteria as the type of task, the size of the task, and the past load on the fog node.

The authors of the study show that the task execution model can be more accurate than the statistical method. This is because the task execution model considers the specific features of the task that can affect its execution time. These features may include the complexity of the task, the skill level of the person performing the task, and the availability of resources.

At the same time, the task execution model can be more difficult to implement than the statistical method. This is due to the fact that the development of a task execution model requires information about the purpose of the distributed system node and its components [9].

A method for calculating the residual resource of fog nodes that takes into account Quality of Service (QoS) requirements is proposed in [10]. The method uses a task execution model to calculate the task execution time, and then uses this execution time to calculate the residual resource.

Implementation of the method consists of the following steps [10]:

QoS Requirements Assessment:
The QoS requirements for the task must first be assessed. QoS requirements may include the following metrics:
- Maximum delay time.
- Maximum error rate.

After calculating the specified metrics, the task execution time is calculated using the task execution model. The task execution model takes into account such criteria as the task type, the task size, and the available resources [11].

Once the task execution time has been calculated, the remaining node element resource can be calculated. The residual resource of a fog node can be calculated by subtracting the resource consumed by the node to perform a task from the total resource of the fog node.

The total node resource requirement for a task is calculated as the product of the node's task execution time and the average resource consumption for one task.

The average resource utilization for task execution can be calculated using a statistical method [12].

However, this method does not account for unforeseen events, such as sudden spikes in load or equipment failures. This can lead to fog node elements being unable to provide the necessary technical specifications for task execution. Additionally, the method does not consider the dynamics of node load, which can change over time. This can lead to the fog node being unable to provide the required quality of task execution if the load and number of tasks change [13].

3. Method Development

To solve the problem of forecasting and calculating the remaining life of elements of fog nodes in distributed information systems of critical infrastructure facilities, it is proposed to use probabilistic forecasting under conditions of resource constraints for regular testing of node elements to assess their technical condition. To do this, we introduce the following restrictions:
We consider a distributed information system with well-defined fuzzy nodes. Each node has its parameters $U_k, k = 1, m$, and $m$ is the total number of nodes. Each parameter has a nominal value $U^0_k$. By observing the dynamics of these parameters, we can evaluate the changes in the processes within the system.

During the operation of a distributed information system with fuzzy nodes, the technical parameters of its elements degrade. This can lead to a deterioration of the parameters of the elements compared to their initial values. This means that at certain points in time $t_i, i = 1, n$ the values of the determining parameters do not increase, i.e. $U_k(t_1) > U_k(t_2) > \cdots > U_k(t_n)$, if $t_1 < t_2 < \cdots < t_n$.

If the values of the parameters $U_k$ are within the specified range of allowable values $S_k = \left(U_k^{(min)} < U_k^0 < U_k^{(max)}\right)$, where $U_k^{(min)}$, $U_k^{(max)}$ are the minimum and maximum allowable values of the parameter, respectively, then the elements function normally [14]. Deviation of the parameter value $U_k$ from its allowable range leads to a parametric failure, which usually does not disable the system, but only degrades its performance. There is also a probability of sudden failures of vague nodes, but in this case there is a parametric failure, which can be recovered from within the framework of recovery work. This means that the distributed information system belongs to the class of highly reliable recoverable systems with a finite number of possible states [15].

We will denote the state that the system can go to at time $t$ as $v_k(t), k = 0, m$. The state of the system, when none of the $m$ defining parameters have failed, will be denoted as $v_0(t)$. The failure of one of the system parameters will be denoted by the symbol $v_1(t)$. The symbol $v_k(t)$ describes the failure of $k$ system parameters, and $v_m(t)$ corresponds to the failure of all $m$-defining parameters. The transition of the system from one state to the nearest neighboring state occurs without a sharp transition of its intermediate state. In other words, the system cannot go from state $v_1(t)$ to state $v_2(t)$ without passing through the state $v_1(t)$.

We will assume that by time $t$, the total cost of maintaining the system in state $v_k(t)$ is $C_k(t), k = 0, m$.

Operating costs $\Delta C_e$ per unit time are the same for each state of the system, and the average restoration costs $\Delta C_n$ are the same for each failed parameter. A critical parameter that fails at time $t$ is restored immediately. During the restoration of the critical parameter, the system costs do not increase. Resources for restoration are allocated at the time of failure of the critical parameters.

The transition of each fuzzy node of the system from one state to another is considered a random event [16].

We will assume that the intensity $\lambda_k$, with which the failure of parameter $k$ of the element $U_k$ occurs, is the same for all defining parameters.

Similar assumptions apply to the parameter $\mu_k$, which characterizes the recovery intensity of the parameter $k$. This approach will allow us to calculate the total failure intensity of the system nodes, which is equal to $\lambda = \sum_{k=1}^{m} \lambda_k$ and $\mu = \sum_{k=1}^{m} \mu_k$, respectively [17].

The introduction of these restrictions requires determining the remaining system resources under conditions of limited resources for the maintenance and repair of fog nodes and their elements.

3.1. Solution

To solve the main problem, it is first necessary to calculate the resource costs for maintaining a separate information system [18] in the state $v_k(t), k = 0, m$.

In $v_0(t)$, the system will be in a failure-free state until time $t + \Delta t$. This means that all nodes and their elements will be operational and functioning normally. The state of the system can be considered as the sum of two events: $A_0$ and $B_1$. These events are incompatible $A_0$.

Event $A_0$ characterizes the state of the system $v_0(t)$ at time $t$. It occurs when there are no failures during $\Delta t$. The probability of this event occurring is described by the equation:

$$P(A_0) = e^{-\lambda \Delta t} \approx 1 - \lambda \Delta t. \quad (1)$$

The resources used to support the event $A_0$ with probability $P(A_0)$ can be calculated using the following formula:

$$C(A_0) = (C_0(t) + \Delta C_e \Delta t)(1 - \lambda \Delta t). \quad (2)$$

If the system is in a state $v_1(t)$ and one of the parameters recovers during the time interval $\Delta t$, the event $B_1$ has occurred [19].
The chance of state $B_1$ occurring is given by the formula:

$$P(B_1) = 1 - e^{-\mu \Delta t} \approx \mu \Delta t. \quad (3)$$

The resource costs for maintaining the system in the state $B_1$ with probability $P(B_1)$ can be determined as follows:

$$C(B_1) = C_1(t) \mu \Delta t. \quad (4)$$

As a result, the system in the state $v_0(t)$ can be considered as a sum of two incompatible events $A_0$ and $B_1$. The equation for the random function of material resource costs $C_0(t)$ on the interval $(t + \Delta t)$ to support this can be expressed as:

$$C_0(t + \Delta t) = C(A_0) + C(B_1) =$$

$$= (C_0(t) + \Delta C_e \Delta t)(1 - \lambda \Delta t) + C_1(t) \mu \Delta t,$$

or like:

$$C_0(t + \Delta t) - C_0(t) =$$

$$= -\lambda C_0(t) \Delta t + \mu C_1(t) + \Delta C_e \Delta t - \lambda \Delta \Delta C_e.$$  

(5)

Divide both sides of equation (6) by $\Delta t$ and take the limit as $\Delta t \to 0$ to obtain the differential equation:

$$\frac{dC_0(t)}{dt} = -\lambda C_0(t) + \mu C_1(t) + \Delta C_e. \quad (7)$$

For system states $v_k(t)$, $k = 0, m - 1$, we write a similar equation in which the state $v_k(t)$ is characterized by the fact that at time $t + \Delta t$ the number of failures of determining parameters is equal to $k$.

Such a state can be considered as a sum of three incompatible events: $A_k, B_{k+1}$, and $C_{k-1}$.

If at the time $t$ the system was in state $v_k(t)$ and during the time $\Delta t$ there were no failures or restorations of any of the parameters, event $A_k$ occurs [20]. The probability of its occurrence is determined by the expression:

$$P(A_k) = e^{-\lambda \Delta t} e^{-\mu \Delta t} \approx$$

$$\approx 1 - (\lambda + \mu) \Delta t. \quad (8)$$

The resource costs for maintaining the system state, which is characterized by event $A_k$ with a probability of occurrence $P(A_k)$ can be determined as follows:

$$C(A_k) = (C_k(t) + \Delta C_e \Delta t)(1 - (\lambda + \mu) \Delta t). \quad (9)$$

The symbol $B_{k+1}$ describes the system at time $t$ and in state $v_{k+1}(t)$ when one parameter has recovered during time $\Delta t$. The probability of its occurrence is determined by the expression:

$$P(B_{k+1}) = 1 - e^{-\mu \Delta t} \approx \mu \Delta t. \quad (10)$$

The resource costs for maintaining the system state described by the event $B_{k+1}$ with probability $P(B_{k+1})$ can be described by the following expression:

$$C(B_{k+1}) = C_{k+1}(t) \mu \Delta t. \quad (11)$$

The next event $C_{k-1}$ is that at time $t$ the system is in state $v_{k-1}(t)$ and during the time $\Delta t$ there is a failure of another parameter. The probability of this event occurring is given by the expression:

$$P(C_{k-1}) = 1 - e^{-\lambda \Delta t} \approx \lambda \Delta t. \quad (12)$$

The resource costs for supporting the event $C_{k-1}$ with probability $P(C_{k-1})$ can be expressed as:

$$C(C_{k-1}) = C_{k-1}(t) \lambda \Delta t. \quad (13)$$

Therefore, the state of the system $v_k(t)$ can be considered as the sum of three incompatible events $A_k, B_{k+1}$ and $C_{k-1}$ and the function of material resource consumption $C_k(t)$ on the interval $(t + \Delta t)$ to support this event can be expressed through:

$$C_k(t + \Delta t) = C(A_k) + C(B_{k+1}) + C(C_{k-1})$$

$$= C_k(t) + \Delta C_e \Delta t(1 - (\lambda + \mu) \Delta t) + C_{k+1}(t) \mu \Delta t + C_{k-1}(t) \lambda \Delta t. \quad (14)$$

Let’s perform a transformation of expression (14) similar to the previous one and we will obtain the differential equation:

$$\frac{dC_k(t)}{dt} = \lambda C_{k+1}(t)$$

$$- (\lambda + \mu) C_k(t) + \mu C_{k+1}(t) + \Delta C_e + \lambda \Delta C_{k+1}. \quad (15)$$

Describing the system state $v_m(t)$, when at time $t + \Delta t$ there will be a failure of all $m$ determining parameters. Such a state can be considered as a sum of two incompatible events: $A_m$ and $C_{m-1}$ [21].

Event $A_m$ characterizes the system in the state $v_m(t)$, at time $t$, when no parameter failures are recorded for a time change of $\Delta t$. The probability of this event is determined by the expression $P(A_m) = e^{-\mu \Delta t} \approx 1 - \mu \Delta t$, and the resource consumption for maintaining the event $A_m$ with probability $P(A_m)$ can be expressed by the equation:

$$C(A_m) = (C_m(t) + \Delta C_e \Delta t)(1 - \mu \Delta t). \quad (16)$$

Event $C_{m-1}$ characterizes the system in the state $v_{m-1}(t)$ at time $t$ and when another
The parameter failure has occurred during the time \( \Delta t \). The probability of this event is determined by the equation
\[
P(C_{m-1}) = 1 - e^{-\lambda \Delta t} = \lambda \Delta t.
\]
The resource costs for supporting the event \( C_{m-1} \) with probability \( P(C_{m-1}) \) can be expressed as:
\[
C(C_{m-1}) = C_{m-1}(t) \lambda \Delta t. \tag{17}
\]
The system in the state \( v_m(t) \) can be represented as the sum of two incompatible events, \( A_m \) and \( C_{m-1} \) [22]. The expression for the function that estimates the costs of material resources \( C_m(t) \) on the interval \( t + \Delta t \) to support this event is given by the following equation:
\[
\frac{\partial C_m(t)}{\partial t} = \lambda C_{m-1}(t) - \mu C_m(t) + C_m(t) + \lambda C_e(t) \Delta \Delta t.
\] (18)

By transforming equation (18), we obtain the following differential equation:
\[
\frac{\partial C_m(t)}{\partial t} = \lambda C_{m-1}(t) - \mu C_m(t) + \lambda C_e(t) \Delta \Delta t.
\] (19)

The costs of material resources to support a system in the state \( v_m(t) \), \( k = 0, m \) can be described by a system of equations involving unknown random functions of costs \( C_k(t), k = 0, m \), which can be solved by transforming it into an operator form [23]:
\[
A(p)X(p) = B^{(c)}(p), \tag{20}
\]
where \( A(p) \) is a three-diagonal matrix of coefficients of a system of equations of dimension \((m + 1) \times (m + 1)\).

\( X(p) \) — vector in which the components \( x_k \) are the mapping functions \( C_k(t), k = 0, m \) [24].

\( B^{(c)}(p) \) — vector of the right-hand sides of the system of equations (20):
\[
\begin{bmatrix}
C_0(p) \\
C_1(p) \\
C_2(p) \\
\vdots \\
C_{m-1}(p) \\
C_m(p)
\end{bmatrix}
= \begin{bmatrix}
\frac{\Delta C_e}{p} \\
\frac{\Delta C_e + \lambda \Delta C_b}{p} \\
\vdots \\
\frac{\Delta C_e + \lambda \Delta C_b}{p}
\end{bmatrix}
\] (21)

\( p \) is a complex variable in the existing half-plane of the mapping functions \( C_k(t), k = 0, m \).

According to [25–28], we will construct an estimate of the vector \( X(p) \) of unknowns of the system of equations (20), using the following procedure for estimating the matrix coefficients of a multidimensional regression analysis model with an arbitrary finite number of regressions.

Let us take the system (20), neglecting the variable \( p \) for the sake of simplicity of the notation and introducing it into the following expression [29]:
\[
\sum_{k=1}^{m+1} A_k X_k = B^{(c)}, \tag{22}
\]
where:
\( A_k \) is the block components of the matrix
\[A = [\|A_1 \| A_2 \| \ldots \| A_k \| \ldots \| A_{m+1} \|].\]
\( X_k \) is the block components of the vector
\[X = [\|X_1 \| X_2 \| \ldots \| X_k \| \ldots \| X_{m+1} \|].\]
\( T \) is the transposition sign.

Using the previously introduced notation, one can find the estimate of \( \hat{X} \) of the vector \( X \) of unknown systems (22) from the recursive relationship:
\[
X_k = C_{k-1} = M_k A_k \prod_{j=1}^{m+1} R_j B^{(c)}, \tag{23}
\]
where
\[M_k = (A_k^T \prod_{j=1}^{k-1} R_j A_k)^{-1},\]
\[R_j = E - A_j M_j A_j^T \prod_{i=1}^{k-1} R_i.\]

Applying the inverse Laplace transform to the estimates (23), we can obtain the expressions of the original random functions of the material costs \( C_k(t) \) for the support of the operation of fog nodes in the distributed information system of a critical infrastructure object \( v_k(t) \), when at the time \( t \) from the total number of \( m \) determining parameters, some of their number \( k = 0, 1, 2, \ldots, m \) may fail [30].

The average material resource expenditure \( \tilde{C}(t) \) is equal to the mathematical expectation of all possible values of the random function \( C_k(t) \):
\[
\tilde{C}(t) = \sum_{k=0}^{m} C_k(t) P_k(t), \tag{24}
\]
where \( P_k(t) \) is the probability of occurrence of the event \( v_k(t) \).

Equation (25) allows us to calculate the average cost of material resources required to maintain fog nodes in each of their possible
states. If the average cost $\bar{C}(t = t)$ at time $t = t$ is close to the additional resources allocated for the operation $C_0$, then we can take $t'$ as the residual resource. To find $t$, we can solve the equation:

$$t = \arg\min_{t \in T} e(C_0 - \bar{C}(t)), \quad (25)$$

where $e(C_0 - \bar{C}(t))$ is the chosen measure of approximation of the quantities $C_0$ and $\bar{C}(t)$:

$T$ is the set of all possible positive values of the variable $t$.

As a result, we will obtain expressions for the quantities $C_k(t), k = 0, m$.

4. Conclusions

This paper examines the efficiency of a distributed information system for a critical infrastructure object.

A distributed information system is a collection of fog nodes and their elements, each with its own functional purpose and physical characteristics. Fog nodes can be physical components such as gateways, switches, routers, and servers, or virtual components such as virtualized switches and virtual machines.

As fog nodes and their elements operate, they consume their technical resources. This can lead to a failure of the critical infrastructure object or its segment.

To reduce the risk of such a failure, this paper proposes a method for calculating the residual technical resource of a fog node and its elements, taking into account their possible failures.

This method can be used to predict the probability of failure of important elements of critical infrastructure objects and to predict the cost of restoring their functionality.

Further research could focus on developing a method for assessing the consequences of malicious attacks on a fog node with low technical resources.

References


